THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 62, No. 7, September 1983
Printed in U.S.A.

Combined Source and Channel Coding for Variable-Bit-Rate Speech Transmission

By D. J. GOODMAN* and C.-E. SUNDBERG†

(Manuscript received January 19, 1983)

Motivated by potential applications to mobile radio, we studied variable-bit-rate speech communication through Gaussian-noise and Rayleigh-fading channels. For convenience we used a constant signaling rate of 32 kb/s and adjusted the source-coding and channel-coding rates in response to changing transmission quality. When the channel quality was good enough, we used all 32 kb/s for speech transmission. When the channel quality was lower, we reduced the source rate to 24 or 16 kb/s and introduced channel coding to control distortion due to transmission errors. We concentrated on specific source and channel codes that could be implemented with bardware of modest complexity. The source code was embedded differential pulse code modulation, which is amenable to variable-bit-rate operation and economical to implement. For error control we introduced punctured convolutional codes and a Viterbi decoder with only 16 states. Although the source/cbannel codec was simple, it offered good performance. Speech quality was at the level of normal telephony when the channel was good; the error-correcting codes extended by up to 13 dB the range of channel signal-to-noise ratios that support adequate quality. Our performance estimates were based on a new analysis of transmission errors in embedded differential pulse code modulation and on computer simulations of speech transmission through fixed and fading channels.

I. INTRODUCTION

1.1 Motivation

Every communication channel has an optimum rate for digital transmission of speech. If the actual rate is below the optimum, speech

* Bell Laboratories. † University of Lund, Sweden.

quality is impaired by unnecessary quantizing distortion. If the rate is higher than optimum, the distortion due to transmission errors is excessive. These observations are embodied in rate-distortion theory, which provides theoretical bounds on performance. Implicitly, they influence practical systems in which design decisions are compromises between goals of quality, bandwidth efficiency, and equipment economy.

When the properties of the communication channel are predictable and invariant, one may explore this compromise and arrive at an acceptable design. However, in many instances (for example, in switched telephony) prior knowledge of the channel is only statistical, and in others (fading microwaves) the channel changes drastically with time. With these channels, performance is characterized statistically by the fraction of users experiencing each possible level of quality or by the fraction of time a single user experiences each quality level.

Both spatial and temporal channel fluctuations are inherent in mobile telephony. Because channel characteristics depend on the position of the mobile unit, all users have, at any time, different channels and, because mobile units are moving, their channels change with time. The conventional approach to this statistical nature of the channel is to set design goals of acceptable quality for a large fraction (say 90 percent) of users or to set "outage" limits, i.e., small fractions of time (say 10 percent) when quality is allowed to fall below a certain specification. In meeting these statistical requirements with a fixed coding and modulation scheme, we have a system in which performance is worse than necessary for most users most of the time. The common deficiency is too much quantizing noise, because 90 percent of the channels are better than the marginal one for which the system is designed. The other 10 percent of the users have too much distortion because of transmission errors.

In this paper we investigate variable-bit-rate transmission, with the information rate adjusted according to the properties of the channel. To facilitate implementation, we maintain a constant signaling rate through the channel and reciprocally adjust the information rate of the source and the amount of channel coding for forward error correction. The result, relative to fixed-rate transmission, is improved grade of service or, for the same grade of service, the ability to serve more users within each geographical area.[1]

We limited the content of this paper to transmission components of a variable-bit-rate system: the source and channel codecs and their performance in two types of channel. Control components, which involve measurements of channel quality and two-way communication between user pairs, require further study.

1.2 Source and channel codes

The most general variable-bit-rate scheme requires control information, an adjustable voice coder/decoder (codec), an adjustable channel codec that provides forward error correction to combat channel impairments, and perhaps an adjustable modem (modulator/demodulator), as indicated in Fig. 1a. In our studies, we have focused on specific source and channel codes that are simple to implement and are especially well suited to variable-bit-rate operation. We do not claim that these codes are optimum, even within a constraint of limited hardware complexity. However, they are strong candidates for practical implementation and this study of their performance provides valuable insights into general principles of variable-bit-rate transmission.

The source codes are embedded Differential Pulse Code Modulation (DPCM),[2] which can be implemented with fixed analog-to-digital and digital-to-analog converters and simple digital control of the number of bits transmitted. The channel codes are convolutional codes, with rates 1/2, 2/3, and 3/4, that provide significant coding gains with maximum likelihood decoding. At rates 2/3 and 3/4, "punctured" code[3] realizations simplify the decoder and lend themselves to variable-bit-

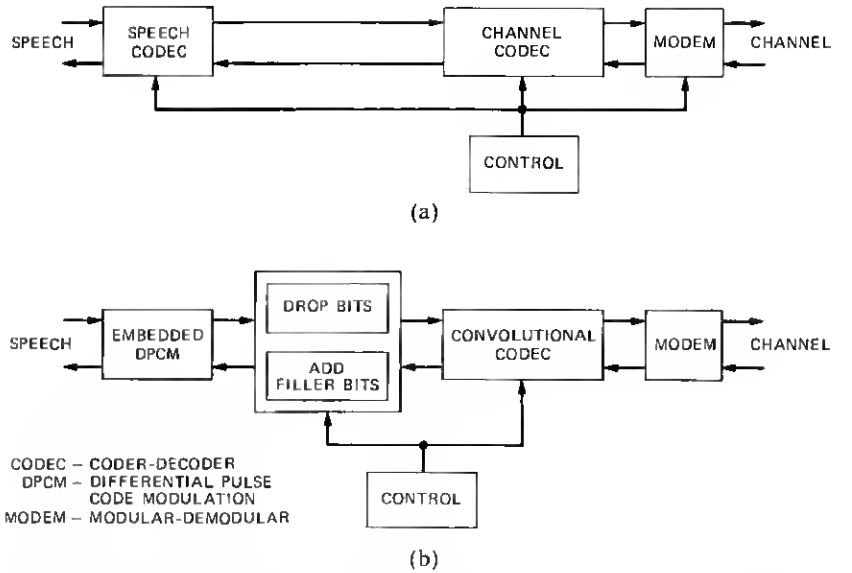

Fig. 1—Elements of a variable-bit-rate system. (a) Adjustable source codec, channel codec, and modem. (b) Fixed embedded source codec and modem. The control of the source rate is simple.

rate operation because the encoder and decoder structures are the same as for the rate 1/2 code. In fact, the punctured codes are rate 1/2 codes with a fraction (1 out of 4 or 2 out of 6) of the channel bits deleted.

In the combined source/channel codec the channel signaling rate is constant; when the channel deteriorates, the least significant speech bits are deleted and the most significant speech bits are protected by a convolutional code. Thus the variable-rate encoding and decoding can be performed, as indicated in Fig. 1b, with a fixed speech codec, a fixed modem, and modest additional hardware relative to fixed-rate operation. Furthermore, the performance characteristics of the embedded source code and the punctured channel code are virtually as good as their conventional counterparts, which require complicated adjustments in a variable-bit-rate environment.

To explore the principle of combining embedded DPCM and punctured convolutional codes we have studied channels with signaling rates of 32 kb/s. We consider four speech transmission formats:

1. All 32 kb/s used for speech transmission
2. 24 kb/s speech transmission, all speech bits protected by a rate 3/4 code
3. 24 kb/s speech transmission, the most significant 2 bits of every 3-bit code word protected by a rate 1/2 code
4. 16 kb/s speech transmission, all speech bits protected by a rate 1/2 code.

These transmission formats, listed in order of increasing resistance to channel impairments, are summarized in Table I, which also presents properties of the convolutional codes discussed in Section III.

Table I—Source and channel code formats, convolutional code properties

	Format 1	Format 2	Format 3	Format 4
1. Source code				
a) bits/sample	4	3	3	2
b) bits/s	32k	24k	24k	16k
c) bits/sample protected	0	3	2	2
2. Channel code				
a) rate	no code	3/4	2/3	1/2
b) $U_0 - U_4$		10101	11001	11101
c) $D_0 - D_4$	encoder, Fig. 4	11111	11011	01011
d) switching pattern		UDDD	UDU	UD
e) free distance, d		4	5	7
f) weight w_d		22	25	4
w_{d+1}	error	0	112	12
w_{d+2}	properties	1687	357	20
w_{d+3}		0	1858	72
w_{d+4}		66964	8406	225

1.3 An example of performance

Figure 2 pertains to phase-shift-keying modulation in a Rayleigh-fading channel. The curves show speech quality, expressed as segmental signal-to-noise ratio (s/n) (Section 2.4.2), as a function of channel s/n. This is defined as the ratio of average energy per channel symbol to noise power per Hertz. Format 1 (32 kb/s speech transmission) is the best choice when the channel is very good (s/n > 22 dB), but it is also the most vulnerable to transmission impairments. By contrast Format 4 (16 kb/s speech), with substantially more quantizing distortion, can pass through very poor channels with no added degradation. Adaptive DPCM at 16 kb/s is highly intelligible, though somewhat fuzzy. Assuming that the threshold of "adequate" performance is 0.5 dB less than the s/n of 16 kb/s error-free transmission, we see in Fig. 2 that Format 4 extends the useful range of channel quality by 13 dB relative to Format 1.

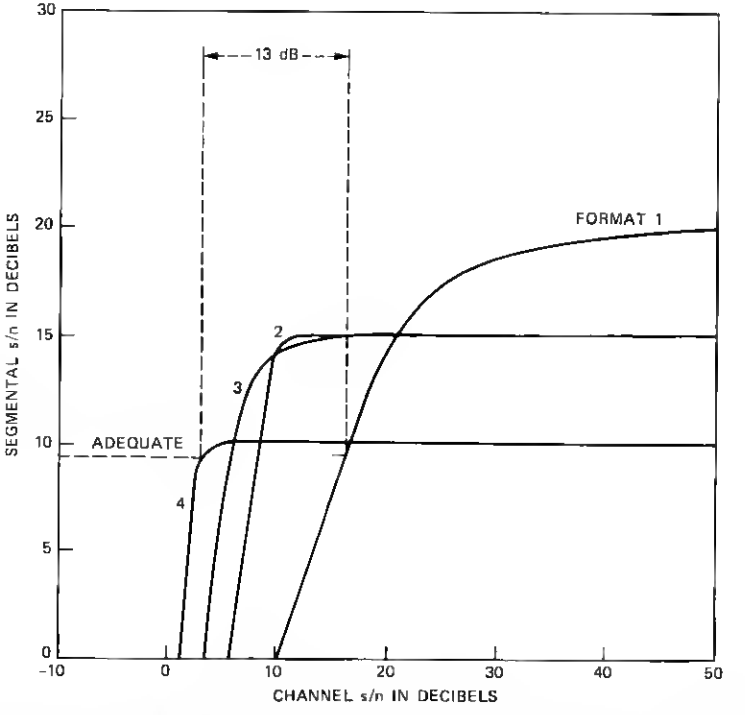

Fig. 2—Performance of the four code formats as a function of the s/n of a Rayleigh-fading channel. If a segmental s/n of 9.5 dB represents "adequate" performance, the variable-rate mechanism extends the threshold of usable channel s/n's from 16 dB (Format 1) to 3 dB (Format 4).

1.4 Organization of this paper

In the following three sections we describe the source and channel codes and present theoretical performance bounds and measures obtained from computer simulations of speech communication over idealized channels. The theory of the effects of transmission errors in embedded DPCM will be published in a separate paper,[4] which extends to embedded DPCM Sundberg and Rydbeck's analysis[5–7] of Pulse Code Modulation (PCM) s/n in noisy channels. Our approacb to combined source coding and channel coding is similar in spirit to the work of Modestino and Daut on image transmission.[8–10] Section V lists several issues to be addressed in evaluating applications of these techniques to mobile radio communication.

II. EMBEDDED DIFFERENTIAL PULSE CODE MODULATION

2.1 Encoder and decoder

Within the bit stream of an embedded code is a slower bit stream that represents the analog signal source with reasonable quality. Embedded coding simplifies variable-bit-rate communication. The analog-to-digital and the digital-to-analog converters always operate at the same high rate. If necessary, the transmission system deletes bits from the source code and adds filler bits prior to decoding. Although PCM is an embedded code, differential PCM, which is more efficient for speech transmission, is not. On the other hand, minor modifications of the conventional DPCM encoder and decoder do produce an embedded code. In the embedded codec, the integrators in the encoder and decoder process a low-bit-rate representation of the error signal. Then, when bits are deleted in transmission, tbe same signal arrives at both integrators and large errors are avoided.

Figure 3 models the embedded DPCM encoder and decoder as the combination of two separate codecs: a "minimal" DPCM codec operating with 2 bits/sample, and a "supplemental" PCM codec, which transmits the quantization error of the minimal encoder. The supplemental codec operates at 2 bits/sample and the combination of minimal and supplemental quantizers can be viewed as a two-stage, successive-approximation realization of a 4-bit/sample quantizer. The 2 bits/sample of the minimal codec are always transmitted, while one or both bits of the supplemental codec can be deleted to reduce the transmission rate from 32 kb/s (4 bits/sample) to 24 or 16 kb/s (3 or 2 bits/sample). The codec functions properly because the encoder and decoder predictors operate on the same signal regardless of the number of bits deleted prior to transmission.

Relative to conventional DPCM, the feedback loop of tbe embedded DPCM encoder operates with reduced resolution when more than 2

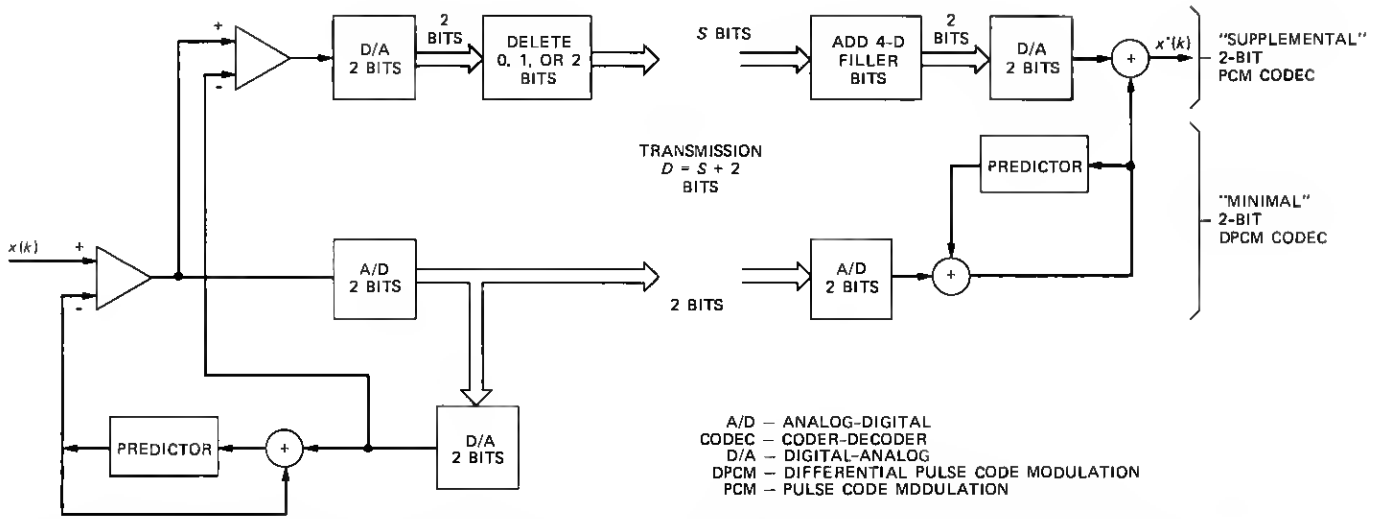

Fig. 3—The embedded DPCM codec modeled as the combination of a "minimal" DPCM codec and a "supplemental" PCM codec. Bits from the supplemental codec can be deleted without severely degrading performance because this deletion has no effect on the ability of the predictor at the decoder to follow the predictor at the encoder.

bits/sample are transmitted. Reference 2 shows that the impact on performance due to this reduced resolution is small, causing no more than a 0.7 dB decrease in s/n in the single-integration codec studied here. (On the other hand, conventional DPCM is unsuited to variable-bit-rate operation. If the transmission system deletes 1 or 2 bits/sample there is a 5.5-dB s/n penalty.)

2.2 Effects of transmission errors

Reference 4 presents a detailed analysis of transmission errors in embedded DPCM. Specialized to a single-integration codec with prediction coefficient a, a key result of that analysis is the formula for the error in the kth sample:

$$x'(k) - x(k) = n_q(k) + e_A(k) + \sum_{i=1}^{\infty} a^i e_M(k - i), \tag{1}$$

in which $x(k)$ and $x'(k)$ are the encoder input and decoder output, respectively. On the right side of (1), $n_q(k)$ is the quantizing distortion of the 2, 3, or 4 transmitted bits/sample, $e_A(k)$ is the noise due to binary errors in all of the transmitted bits, and $e_M(k)$ is the noise due to binary errors in the 2 bits/sample of the minimal quantizer. Equation (1) shows that the errors in these two bits are amplified by the integrator in the decoder.

2.3 Audio signal-to-noise ratio

Our analytic tools[4] enable us to compute the mean-square value of (1) under the condition that the quantizer is in a granular state (output and input differ by no more than half a quantizing step). The derived signal-to-noise ratio, therefore, fails to reflect overload distortion. This apparent deficiency does not limit the value of the granular mean-square error as a predictor of speech quality. Subjective tests show that granular noise and overload are quite different perceptually and that adding their squares is often misleading.

A similar issue is raised by the addition in (1) of a granular-noise term (n_q) to e_A and e_M, which result from transmission errors. Is the mean-square sum a meaningful measure of signal quality? We are optimistic that it is because, contrary to overload, which is signal-dependent, granular noise and noise due to transmission errors are essentially uncorrelated with tbe input. With this rationale we have derived signal-to-noise ratio formulas of the form[4]

$$\text{s/n} = \frac{E[x(k)]^2}{E[x'(k) - x(k)]^2} = \frac{C}{\sigma_q^2 + \sigma_{qt}^2}, \tag{2}$$

in which C depends on the codec configuration and the source statistics, σ_q^2 is the quantizing noise power, and σ_{qt}^2 is the distortion due to

transmission errors, including the effects of correlation between quantizing noise and noise due to transmission errors.

With single-integration DPCM and a 2-bit minimal quantizer

$$C = \frac{1 - a^2L^2/48}{L^2(1 - 2a\rho + a^2)} \tag{3}$$

and

$$\sigma_q^2 \approx 2^{-2D}/3, \tag{4}$$

where D (2, 3, or 4) is the number of bits transmitted, a is the predictor coefficient of the DPCM feedback loop, ρ is the adjacent-sample autocorrelation coefficient of the encoder input signal, $x(k)$, and L is the quantizer load factor: the ratio of overload point to rms quantizer input.

The other quantity in (2), σ_{qt}^2, depends on the binary representation of quantizer output signals and on the transmission conditions including the modulator and demodulator, the channel, and the encoder and decoder (if they are included) for forward error correction. Table II contains the formulas for σ_{qt}^2 for the four transmission formats studied in this paper. The formulas pertain to a sign-magnitude binary representation and a quantizer with "3.16-sigma loading", i.e., the quantizer overload point is $\sqrt{10}$ times the rms value of the quantizer input.

Table II—Formulas for computing audio signal-to-noise ratio

General formula

$$s/n = \frac{(1 - a^2L^2/48)}{L^2(1 - 2a\rho + a^2)} \cdot \frac{1}{(\sigma_{qt}^2 + 2^{-2D}/3)}$$

Coder parameters

Predictor coefficient $a = 0.85$

Quantizer load factor $L = \sqrt{10}$

Input signal

Correlation coefficient $\rho = 0.85$

Transmission effects, sign-magnitude representation

P: binary error probability of unprotected bits

P_e: binary error probability of convolutional code

Format 1: $D = 4$ (no channel coding)

$$\sigma_{qt}^2 = 0.723P + 0.271P^2 + \frac{a^2}{1 - a^2}(0.725P + 0.275P^2)$$

Format 2: $D = 3$ (rate 3/4 code)

$$\sigma_{qt}^2 = 0.25P_e\left(3.37 + \frac{a^2}{1 - a^2}1.73\right)$$

Format 3: $D = 3$ (rate 2/3 code)

$$\sigma_{qt}^2 = 0.5P_e\left(1.52 + \frac{a^2}{1 - a^2}1.73\right) + 0.0625P$$

Format 4: $D = 2$ (rate 1/2 code)

$$\sigma_{qt}^2 = P_e\left(1.56 + \frac{a^2}{1 - a^2}1.73\right)$$

2.4 Speech transmission

2.4.1 Adaptive DPCM

Practical DPCM codecs have time-varying step sizes to accommodate the wide dynamic range of speech sounds. In this study, we have used a robust adaptive quantizer[11] in which the step size of the encoder changes with each sample according to the rule

$$\delta_{k+1} = M\delta_k^{\beta}, \tag{5}$$

where δ_k is the step size used to encode the kth sample, M is a multiplicative factor that depends on the most recent encoder output, and β is a leakage factor that belps the codec work properly in the presence of transmission errors.

At the decoder the step size is δ_k' with

$$\delta_{k+1}' = M'(\delta_k')^{\beta}, \tag{6}$$

where M' depends on the received code word. It can differ from M if there has been an error in transmitting the kth code word.

To make the adaptive quantizer work with embedded DPCM, we restrict M to one of two possible values:

$$M = M_1 < 1$$

if the input is in the lower half of the quantizer range, and

$$M = M_2 > 1$$

otherwise. Then with four-bit code words represented in a sign-magnitude format, M and M' depend only on the most significant magnitude bit. As a consequence the decoder step size can track the encoder step size when one or both of the two least significant bits are deleted.

In the simulation studies reported in this paper the adaptive quantizer constants were $M_1 = 0.85$, $M_2 = 1.5$, and $\beta = 0.98$. We selected these values because they offer a good compromise between performance over an ideal channel and tolerance of transmission errors.

2.4.2 Segmental signal-to-noise ratio

In addition to the limitations discussed in Section 2.3, the s/n formulas of Table II are of limited value in evaluating adaptive DPCM coding of speech because they do not account for errors due to different step sizes at the transmitter and receiver. On the other hand, segmental s/n is a quality measure that is reasonably well correlated with listener reports of the quality of speech transmitted by adaptive DPCM over noisy channels. It is defined as the average of the s/n measured in decibels over short segments of the signal, and it can be calculated in a straightforward manner in computer simulations of speech trans-

mission.[12] In our simulations we measure s/n over 16-ms segments (128 samples) and exclude from the average segments detected as silent (power at least 55 dB below the saturation level of the quantizer). We also limit the s/n of each segment to the range −10 to 80 dB.

III. CONVOLUTIONAL CODING AND DECODING, PSK MODULATION

Our design, analysis, and implementation of the rate 3/4, 2/3, and 1/2 coders and decoders for the most part follow Chapter 6 of Ref. 13. The punctured codes[3] (rate 1/2 codes with a fraction of the channel symbols deleted) are especially well suited to variable-bit-rate operation because the encoder and decoder retain their basic structures as the code format changes; the things that do change are a set of constants (code generators) and the switching patterns that govern the output of channel symbols at the encoder and the input of channel symbols at the decoder.

3.1 Code configuration

In our calculations and simulations, we use codes generated according to Fig. 4. With a five-stage shift register at the encoder, the codes have 4-bit memory (16 unique nodes in the code trellis). This number provides an attractive compromise between performance (enhanced by long memory) and decoder simplicity (short memory). When the rate 1/2 code is employed, the DPCM signal arrives at 16 kb/s, and for each input bit, the switch at the encoder gets one output bit from the top branch of Fig. 4 and one output bit from the bottom branch. When the rate is 2/3, the two most significant bits in each 3-bit DPCM code word stimulate three output bits from Fig. 4. The first of the two DPCM bits results in two output bits—one from the upper branch of Fig. 4 and one from the lower branch, as in the rate 1/2 code. When the second DPCM bit arrives, the encoder releases one bit from the upper branch. Along with these three bits from the convolutional encoder the least significant DPCM bit is transmitted without protection. When the rate is 3/4, the first member of every block of three input bits results in two output bits (from upper and lower branch of Fig. 4); the remaining two input bits in the block stimulate one output bit each (both from the lower branch). The code generators and switching patterns are listed in Table I.

The receiver is a maximum-likelihood decoder, implementing the Viterbi Algorithm. A signal processor produces a measurement (between 0 and 1) for each received bit. This measurement is near 0 if there is a high likelihood that the bit was transmitted as zero and near 1 if there is a high likelihood that one was transmitted. With each block of two (rate 1/2), three (rate 2/3), or four (rate 3/4) received bits, the Viterbi decoder accepts combined measurements of the first

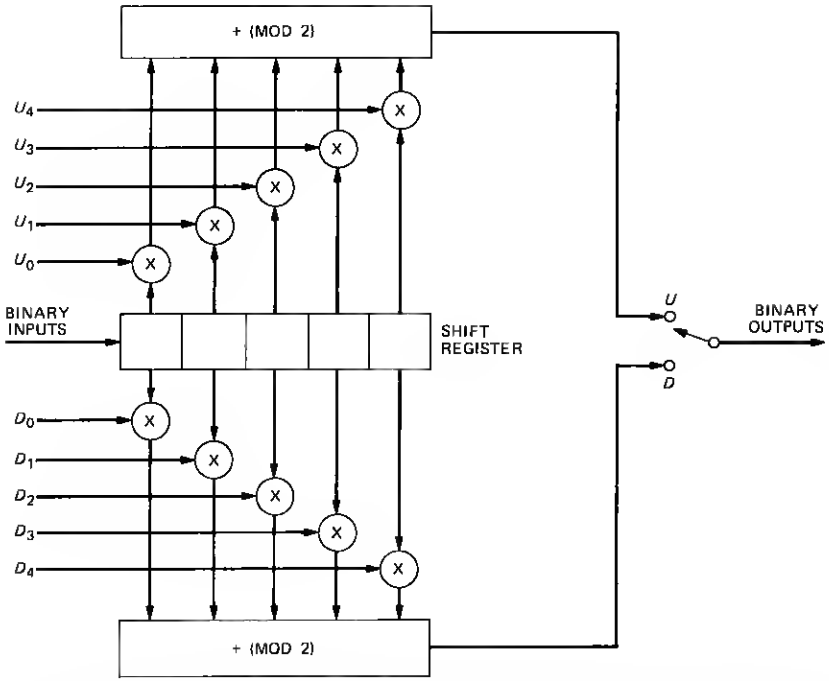

Fig. 4—Structure of the convolutional encoder. U_i and D_i are binary coefficients. For rate 1/2 code there are two output bits (U and D) for each input bit. For rate 2/3 there are two outputs (U and D) for one input and one output (U) for the next input. For rate 3/4 there are two outputs (U and D) for the first input and one output (D) for each of the next two input bits.

two bits and updates the metrics and path memories associated with the 16 possible decoder states. The decoder then releases one bit corresponding to an input bit delayed by the length of the decoder path memory. (In the simulations reported here the path memory was 30 bits for all codes.) If the rate is 2/3 or 3/4, the measurement of each remaining bit in the block causes the decoder to update all metrics and path memories and to release one additional bit.

3.2 PSK modulation, likelihood measurements

To study the performance of variable-bit-rate transmission with embedded DPCM and convolutional codes we have confined our attention to Binary Phase-Shift Keying (BPSK) modulation with coherent detection. In a Gaussian-noise channel the demodulator output for each channel symbol is either

$$r = -A + n \quad \text{or} \quad r = A + n,$$

depending on whether "zero" or "one" was transmitted. The constant,

A, is the received sinewave amplitude and the noise, n, is a sample of a zero-mean normal random variable with variance σ^2. The symbol s/n is $\rho = A^2/2\sigma^2$. In our computer simulations, the likelihood measurement for each symbol is $m = 0.5 + r/\text{const}$, where the constant is large enough to ensure $0 < m < 1$. (With m effectively unquantized, as in our simulations, the constant is otherwise arbitrary. With m quantized to a few bits the constant is a compromise between dynamic range and resolution.)

To update the state metrics of the Viterbi decoder we add m or $1 - m$, depending on whether the relevant branch of the code tree is associated with zero or one transmitted. For a constant-amplitude, white-Gaussian-noise channel, this procedure produces a precise likelihood metric. In a fading channel, which is of principal interest in mobile radio, this metric is only an approximation to the likelihood function.[14] A true likelihood metric involves elaborate computations that would probably be prohibitively complicated in practice.

3.3 Probability of error

The theory of convolutional coding provides upper (union) bounds on bit-error probability as a function of s/n. These bounds are infinite series that are truncated for the purpose of numerical computation. In our calculations we have used sums of five terms to estimate the error probabilities of the convolutional codes:

$$P_e = \frac{1}{b} \sum_{n=d}^{d+4} w_n f_{n,C}(\rho), \tag{7}$$

where d is the free distance of the code; b is the number of source bits per code word (1, 2, 3 for rate 1/2, 2/3, 3/4); and $f_{n,C}(\rho)$ is the probability that an incorrect path, n bits removed from the correct path, has a lower metric than the correct path. This probability is a function of the type of channel, ($C = g$ for a Gaussian channel, $C = R$ for Rayleigh fading), the modulation technique, and the s/n ρ.

Table I contains the free distances, d, and the weights, w_n, of the three convolutional codes we have studied. For BPSK with coherent detection we have the precise formula for a nonfading channel,

$$f_{n,g}(\rho) = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2n\rho}}^{\infty} \exp\left[-\frac{x^2}{2}\right] dx. \tag{8}$$

For independent Rayleigh fading signals in white Gaussian noise, we do not have a precise formula for $f_{n,R}(\rho)$. Instead we use the upper bound[14]

$$f_{n,R}(\rho) < \left[\exp(\beta/\rho)\, \frac{\sqrt{1 + 2/\beta} - 1}{\sqrt{1 + 2/\beta} + 1}\right]^n, \tag{9}$$

$$\beta = \sqrt{\rho^2 + 1} - 1. \tag{10}$$

Figures 5 and 6 show estimates of binary error rates computed from (7) through (10) and the results of computer simulations of random data transmitted through a convolutional coder, a random channel, and a Viterbi decoder with path memory 30 bits. In Fig. 5 the simulations and calculations for the nonfading channel agree very closely. On the other hand Fig. 6 suggests that the bound in (9) is quite loose over the range of conditions of interest to us here; the estimated channel s/n corresponding to a given error rate is 3 to 4 dB lower in the simulations than in the computed curves. Figures 5 and

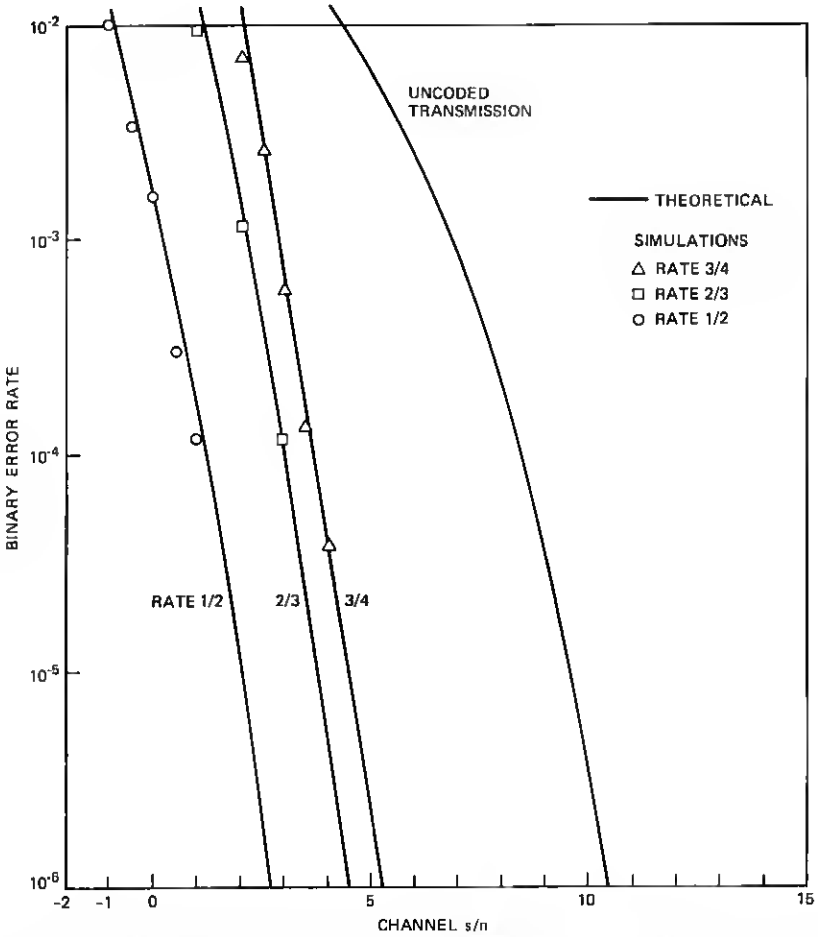

Fig. 5—Binary error rates of convolutional codes and uncoded Phase-Shift Keying (PSK) in a nonfading channel. The theoretical curves and the results of computer simulations are in close agreement.

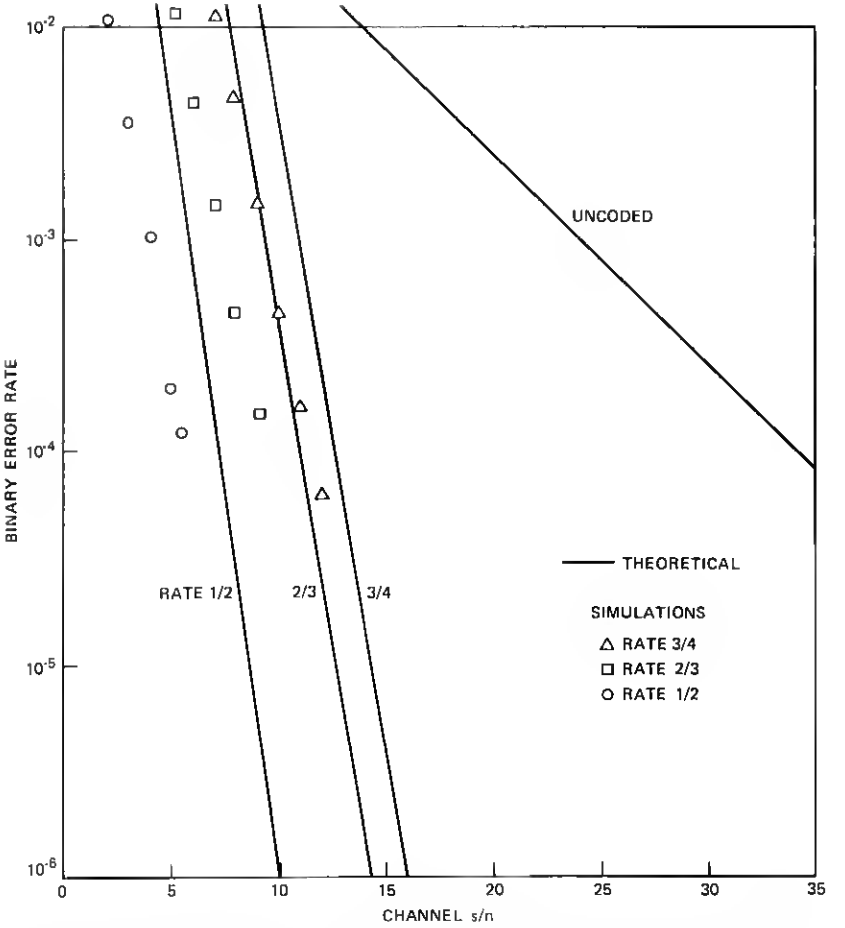

Fig. 6—Binary error rates of convolutional codes and uncoded PSK in a Rayleigh-fading channel. The theoretical curves, based on upper bounds on path error probabilities, are loose over the range of error rates (10^{-4} to 10^{-2}) of interest to us here.

6 also show P, the binary error rates for uncoded transmission, which are $P = f_{1,g}(\rho)$ in (8) for the nonfading channel and

$$P = \frac{1}{2}\left[1 - \sqrt{\frac{\rho}{1+\rho}}\right] \tag{11}$$

with Rayleigh fading.[15]

Note that the independent variable in Figs. 5 and 6 is the s/n of *channel symbols* rather the s/n of *information bits*, which is often plotted in comparisons of coding schemes for digital data transmission.

For our purpose the channel s/n is an appropriate measure because it remains constant as we change code formats.

IV. PERFORMANCE

To evaluate the potential effectiveness of variable-bit-rate transmission we have computed audio s/n as a function of channel s/n using the formulas in Table II and (7) through (10). Figure 7 applies to a nonfading Gaussian channel and Fig. 8 applies to independent Rayleigh fading signals (the s/n's of all channel symbols are mutually independent) in Gaussian noise. All curves pertain to a sign-magnitude representation of source symbols, which is generally more tolerant of transmission errors than the natural-binary representation.

We assume that an audio s/n within 0.5 dB of the s/n of error-free 16 kb/s transmission provides "adequate" voice quality; therefore, in Fig. 7 we estimate that relative to conventional 32 kb/s transmission, variable-bit-rate operation extends the range of useful channel conditions by 4.3 dB in a nonfading channel. (Without channel coding

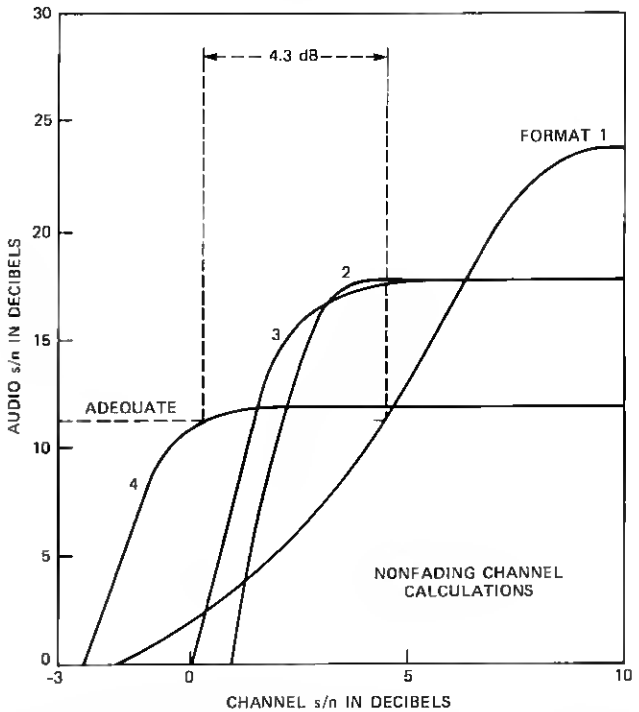

Fig. 7—Calculated performance of the four code formats as a function of the s/n of a nonfading channel. Convolutional coding extends the range of channel s/n's that offer adequate communication by 4.3 dB.

voice quality is inadequate when the channel s/n is less than 4.5 dB. With coding this threshold is 0.2 dB.) Figure 8 suggests that with Rayleigh fading the coding gain is 9.2 dB.

The theoretical formulas, though simple, rely on many bounds, unverified assumptions, and idealizations of a practical communications environment. To obtain more realistic estimates of performance we have resorted to computer simulations of adaptive DPCM transmission of a 2.5-second speech sample. The simulated Viterbi decoder has a path memory of 30 bits. Performance measurements, with speech quality measured as segmental s/n (see Section 2.4.2), are shown in Fig. 9 for a nonfading channel and in Fig. 2 for the Rayleigh-fading channel.

We see that the approximations in Figs. 7 and 8 provide the same qualitative information as the speech simulations in Figs. 9 and 2. For the nonfading channel, simulations and theory lead to nearly equal estimates (4.3 dB and 4.7 dB) of coding gain. In the case of Rayleigh fading the theory underestimates this gain (by 3.8 dB) relative to

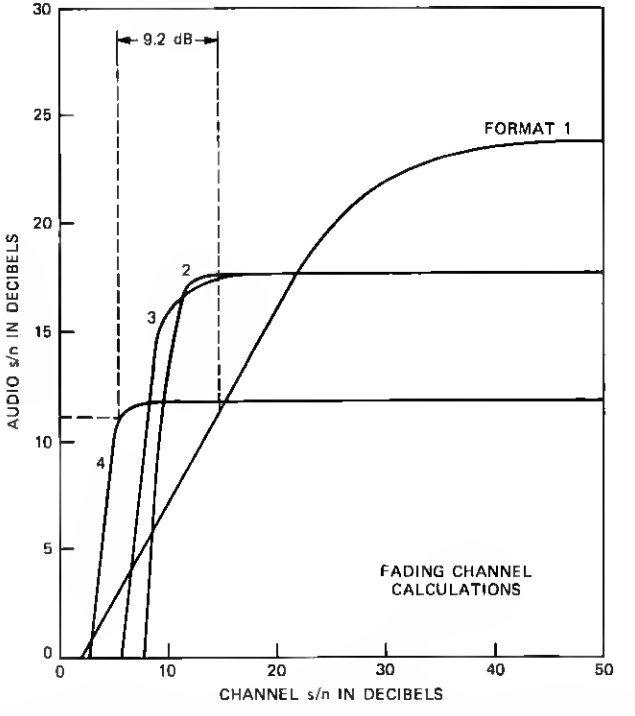

Fig. 8—Calculated performance of the four code formats in a fading channel. The estimate of 9.2-dB coding gain is a lower bound because the computed error probabilities of the Viterbi decoder are loose upper bounds.

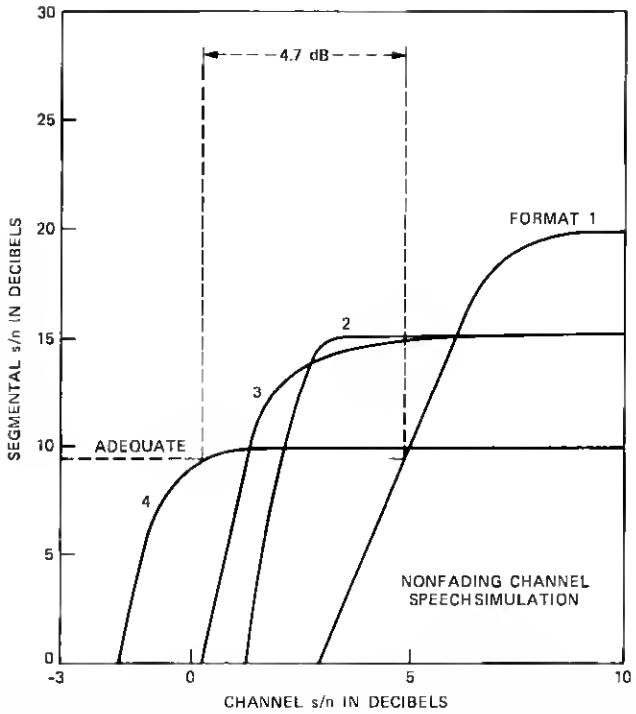

Fig. 9—Performance of the four code formats in simulated speech transmission over a nonfading channel. Coding lowers by 4.7 dB the channel s/n required for adequate quality.

simulation. This discrepancy is largely due to the loose upper bound on the binary error rate of the Viterbi decoder in a fading channel. (See Fig. 6 and Section 3.3.)

All of the performance curves show that transmission Format 2 (24 kb/s speech, all bits protected) is of very limited value. When the coded bits are essentially error free, Format 2 is only slightly better than Format 3 (24 kb/s, 2 of 3 bits/sample protected). In this condition, the transmission noise of Format 3 is due to errors in the third, uncoded, bit, which have a small effect on the quality of embedded DPCM because the errors are not enhanced by the receiver integrator. (In nonembedded DPCM Format 2 would be more effective than Format 3 over a wide range of channel conditions.[4]) In a practical application, therefore, Format 2 would be omitted and the system would switch among Formats 1, 3, and 4.

V. CONCLUSIONS: APPLICATIONS TO MOBILE RADIO

We have described the key elements of a combined source/channel

codec and evaluated its ability to communicate speech over idealized channels. While this approach has other possible applications, we were motivated to study its potential for enhancing mobile radiotelephony. A detailed study of the codec in a mobile-radio environment is currently under investigation. We conclude this paper by reviewing the features of the scheme and listing important issues to be addressed in assessing its value in a mobile-radio context.

A principal advantage of the source codes and channel codes is their simplicity. The embedded DPCM coder and decoder can be realized on a single-chip microcomputer[16] and the Viterbi decoder, with only 16 states, is within the state-of-the-art of special-purpose integrated circuits.[17] Because the channel signaling rate is constant, no special modem is required. Notwithstanding this simplicity, the speech quality for good channels is comparable to that of conventional telephony and the extension of the useful range of channel conditions (4.5 dB in a nonfading channel, 13 dB with independent Rayleigh fading) is substantial.

To extend this idealized study to the context of variable-bit-rate communication over mobile-radio channels, we are led to study the:

1. Control mechanism for altering code formats
2. Channel characteristics, including the temporal nature of the Rayleigh fading and the geographical distribution of signal-to-interference ratio in a mobile-radio service area
3. Alternative modulation techniques to binary-phase-shift keying
4. Possibility of diversity reception.

REFERENCES

1. D. J. Goodman and C.-E. Sundberg, "Quality of Service and Bandwidth Efficiency of Cellular Mobile Radio with Variable-Bit-Rate Speech Transmission," Trans. Veh. Technol., *VT-32*, No. 3 (August 1983).
2. D. J. Goodman, "Embedded DPCM for Variable Bit Rate Transmission," IEEE Trans. Commun., *COM-28*, No. 7 (July 1980), pp. 1040–6.
3. J. B. Cain, G. C. Clark, and J. M. Geist, "Punctured Convolutional Codes of Rate $(n-1)/n$ and Simplified Maximum Likelihood Decoding," IEEE Trans. Inf. Theory, *IT-25* (January 1979), pp. 97–100.
4. D. J. Goodman and C.-E. Sundberg, "Transmission Errors and Forward Error Correction in Embedded Differential Pulse Code Modulation," B.S.T.J., *62*, No. 9, Part 1 (November 1983).
5. N. Rydbeck and C.-E. Sundberg, "Analysis of Digital Errors in Non-Linear PCM Systems," IEEE Trans. Commun., *COM-24*, No. 1 (January 1976), pp. 59–65.
6. C.-E. Sundberg and N. Rydbeck, "Pulse Code Modulation with Error-Correcting Codes for TDMA Satellite Communication Systems," Ericsson Technics, *32*, No. 1 (1976), pp. 1–56.
7. N. Rydbeck and C.-E. Sundberg, "PCM/TDMA Satellite Communication Systems with Error Correcting and Error Detecting Codes," Ericsson Technics, *32*, No. 3 (1976), pp. 195–247.
8. J. W. Modestino and D. G. Daut, "Combined Source and Channel Coding of Images," IEEE Trans. Commun., *COM-27*, No. 11 (November 1979), pp. 1644–59.
9. J. W. Modestino, D. G. Daut, and A. L. Vickers, "Combined Source-Channel Coding of Images Using the Block Cosine Transform," IEEE Trans. Commun., *COM-29*, No. 9 (September 1981), pp. 1261–74.

10. D. G. Daut and J. W. Modestino, "Two-Dimensional DPCM Image Transmission Over Fading Channels," IEEE Trans. Commun., *62*, No. 9, Part 1 (November 1983).
11. D. J. Goodman and R. W. Wilkinson, "A Robust Adaptive Quantizer," IEEE Trans. Commun., *COM-23*, No. 11 (November 1975), pp. 1362–5.
12. C. Scagliola, "Evaluation of Adaptive Speech Coders Under Noisy Channel Conditions," B.S.T.J., *58*, No. 6, Part 2 (July–August 1979), pp. 1369–94.
13. G. C. Clark and J. B. Cain, *Error-Correction Coding for Digital Communications*, New York: Plenum Press, 1981, Chapter 6, pp. 227–66.
14. J. Hagenauer, "Viterbi Decoding of Convolutional Codes for Fading and Burst Channels," Proc. 1980 Int. Zurich Seminar (March 1980) pp. G2.1–7.
15. M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*, New York: McGraw Hill, 1966, p. 407.
16. J. R. Boddie, et al., "Adaptive Differential Pulse-Code-Modulation Coding," B.S.T.J., *60*, No. 7, Part 2 (September 1981), pp. 1547–61.
17. R. M. Orndorff, et al., "Viterbi Decoder VLSI Integrated Circuit for Bit Error Correction," Nat. Telecommun. Conf., New Orleans, Louisiana, December 1981, Conference Record, pp. E1.7.1–4.

AUTHORS

David J. Goodman, B.E.E., 1960, Rensselaer Polytechnic Institute; M.E.E., 1962, New York University; Ph.D. (Electrical Engineering), 1967, Imperial College, London; Bell Laboratories, 1967—. Mr. Goodman has studied various aspects of digital communications including analog-to-digital conversion, digital signal processing, subjective assessment of voiceband codecs, and spread spectrum modulation for mobile radio. He is Head, Communications Methods Research Department. In 1974 and 1975 he was a Senior Research Fellow and in 1983 a Visiting Professor at Imperial College, London, England. Member, IEEE.

Carl-Erik W. Sundberg, M.S.E.E., 1966, and Dr. Techn., 1975, Lund Institute of Technology, University of Lund, Sweden; Bell Laboratories, 1981–1982. Mr. Sundberg is an Associate Professor in the Department of Telecommunication Theory, University of Lund, and a consultant in his field. He is Director of the consulting company SUNCOM, Lund. During 1976 he was with the European Space Research and Technology Centre (ESTEC), Noordwijk, The Netherlands, as an ESA Research Fellow. He has been a Consulting Scientist at LM Ericsson and SAAB-SCANIA, Sweden, and at Bell Laboratories. His research interests include source coding, channel coding (especially decoding techniques), digital modulation methods, fault-tolerant systems, digital mobile radio systems, spread spectrum systems, and digital satellite communication systems. He has published a large number of papers in these areas during the last few years. Senior Member, IEEE; member, SER, Sveriges Elektroingenjörers Riksförening.